

University of Groningen

HST large-field weak lensing analysis of MS 2053-04

Hoekstra, Henk; Franx, Marijn; Kuijken, Konrad; van Dokkum, Pieter G.

Published in:
Monthly Notices of the Royal Astronomical Society

DOI:
[10.1046/j.1365-8711.2002.05479.x](https://doi.org/10.1046/j.1365-8711.2002.05479.x)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2002

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Hoekstra, H., Franx, M., Kuijken, K., & van Dokkum, P. G. (2002). HST large-field weak lensing analysis of MS 2053-04: study of the mass distribution and mass-to-light ratio of X-ray luminous clusters at $0.22 < z < 0.83$. *Monthly Notices of the Royal Astronomical Society*, 333(4), 911-922. <https://doi.org/10.1046/j.1365-8711.2002.05479.x>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

***HST* large-field weak lensing analysis of MS 2053–04: study of the mass distribution and mass-to-light ratio of X-ray luminous clusters at $0.22 < z < 0.83$**

Henk Hoekstra,^{1,2,3★} Marijn Franx,⁴ Konrad Kuijken³ and Pieter G. van Dokkum^{5†}

¹*CITA, University of Toronto, 60 St. George Street, Toronto, M5S 3H8, Canada*

²*Department of Astronomy, University of Toronto, 60 St. George Street, Toronto, M5S 3H8, Canada*

³*Kapteyn Astronomical Institute, University of Groningen, Postbus 800, 9700 AV Groningen, the Netherlands*

⁴*Leiden Observatory, P. O. Box 9513, 2300 RA Leiden, the Netherlands*

⁵*California Institute of Technology MS 105-24, Pasadena, CA 91125, USA*

Accepted 2002 March 3. Received 2002 February 14; in original form 2001 October 3

ABSTRACT

We have detected the weak lensing signal induced by the cluster of galaxies MS 2053–04 ($z = 0.58$) from a two-colour mosaic of six *Hubble Space Telescope* (*HST*) WFPC2 images. The best-fitting singular isothermal sphere model to the observed tangential distortion yields an Einstein radius $r_E = 6.2 \pm 1.8$ arcsec, which corresponds to a velocity dispersion of 886^{+121}_{-139} km s^{−1} ($\Omega_m = 0.3$, $\Omega_\Lambda = 0.0$). This result is in good agreement with the observed velocity dispersion of 817 ± 80 km s^{−1} from cluster members. The observed average rest-frame mass-to-light ratio within a $1 h_{50}^{-1}$ Mpc radius aperture is $184 \pm 56 h_{50} M_\odot/L_{B\odot}$. After correction for luminosity evolution to $z = 0$, this value changes to $291 \pm 89 \pm 19 h_{50} M_\odot/L_{B\odot}$ (where the first error indicates the statistical uncertainty in the measurement of the mass-to-light ratio, and the second error is due to the uncertainty in luminosity evolution).

MS 2053 is the third cluster we have studied using mosaics of deep WFPC2 images. For all three clusters we find good agreement between dynamical and weak lensing velocity dispersions, in contrast to weak lensing studies based on single WFPC2 pointings on cluster cores. This result demonstrates the importance of wide-field data. We have compared the ensemble-averaged cluster profile of the clusters in our sample with the predicted Navarro, Frenk & White (NFW) profile, and find that an NFW profile can fit the observed lensing signal well. The best-fitting concentration parameter is found to be $0.79^{+0.44}_{-0.15}$ (68 per cent confidence) times the predicted value from an open cold dark matter (CDM) model. The observed mass-to-light ratios of the clusters in our sample evolve with redshift, and are inconsistent with a constant, non-evolving, mass-to-light ratio at the 99 per cent confidence level. The evolution is consistent with the results derived from the evolution of the fundamental plane of early-type galaxies. The resulting average mass-to-light ratio for massive clusters at $z = 0$ is found to be $239 \pm 18 \pm 9 h_{50} M_\odot/L_{B\odot}$.

Key words: gravitational lensing – galaxies: clusters: individual: MS 2053–04 – galaxies: fundamental parameters – X-rays: galaxies: clusters.

1 INTRODUCTION

Observations of high-redshift clusters are valuable to test our current understanding of structure formation on cosmological scales (e.g. Eke, Cole & Frenk 1996; Bahcall & Fan 1998). In

particular, reliable mass estimates of these systems are important, as they provide strong constraints on cosmological models.

The small, systematic, distortion in the shapes of background sources induced by massive structures, known as weak gravitational lensing, has proven to be a powerful method to measure the masses of clusters of galaxies (for an extensive review see Mellier 1999). The weak lensing effect allows one to reconstruct the projected surface mass density (e.g. Kaiser & Squires 1993) or

★E-mail: hoekstra@cita.utoronto.ca

†Hubble Fellow.

measure the mass, without having to rely on assumptions about the state or nature of the deflecting matter. However, for an accurate mass estimate, high number densities of background galaxies are needed, as well as a good estimate of their redshift distribution.

Lensing studies of high-redshift clusters ($z > 0.5$) are difficult because the lensing signal is low and most of the signal comes from small, faint sources. These sources typically have sizes that are comparable to the size of the point spread function (PSF) in ground-based images. To extract the lensing signal from such observations, large corrections are required. For these studies *Hubble Space Telescope* (HST) observations have great advantage over ground-based observations because the background sources are much better resolved, resulting in a well-calibrated weak lensing signal. However, the mass distribution of clusters with redshifts $z < 0.4$ can be studied well from the ground as much of the lensing signal comes from source galaxies at redshifts $z \sim 0.8$, whose shapes can be measured well in ground-based observations.

In this paper we present the results of our weak lensing analysis of the $z = 0.58$ cluster of galaxies MS 2053–04. It is the third cluster for which we have studied the mass distribution based on a deep two-colour mosaic of WFPC2 images. The other two clusters that have been studied this way are CI 1358+62 ($z = 0.33$) (Hoekstra et al. 1998, hereafter HFKS) and MS 1054–03 ($z = 0.83$) (Hoekstra, Franx & Kuijken 2000, hereafter HFK). All three clusters have been selected on the basis of their strong X-ray emission.

MS 2053 was detected in the *Einstein* Medium Sensitivity Survey (EMSS) (Gioia & Luppino 1994). It is one of the few $z > 0.5$ clusters found in this survey, and of these high-redshift clusters it has the lowest X-ray luminosity. Its X-ray luminosity¹ is $L_X(2–10 \text{ keV}) = (7.9 \pm 0.7) \times 10^{44} h_{50}^{-2} \text{ erg s}^{-1}$ (Henry 2000). The X-ray temperature measured by *BeppoSAX* is $kT = 6.7_{-2.3}^{+6.8} \text{ keV}$ (Della Ceca et al. 2000). A more accurate temperature of $kT = 8.1_{-2.2}^{+3.7} \text{ keV}$ has been determined from *ASCA* observations (Henry 2000).

Luppino & Gioia (1992) discovered a gravitationally lensed arc in deep images of MS 2053. The arc is located approximately 15 arcsec from the Brightest Cluster Galaxy (BCG). Its redshift is still unknown. The cluster mass distribution has been studied previously through weak lensing by Clowe (1998) based on deep ground-based images.

We first present the results of the weak lensing analysis of MS 2053. In Section 2 we briefly discuss the data, and in Section 3 the object analysis is described. The cluster light distribution is examined in Section 4. In Section 5 we present the weak lensing signal and the reconstruction of the projected surface mass density. The mass and mass-to-light ratio inferred from our analysis are presented in Section 6. In Section 7 we present the combined results of a sample of four clusters that have been studied through weak lensing. We compare the weak lensing mass estimates to dynamical estimates. We also study the average mass profile of the clusters, as well as their mass-to-light ratios.

2 DATA

To study the cluster MS 2053 we use a mosaic of WFPC2 images taken with the *HST*. Fig. 1 shows the layout of the mosaic constructed from the six pointings of the telescope. The cluster has

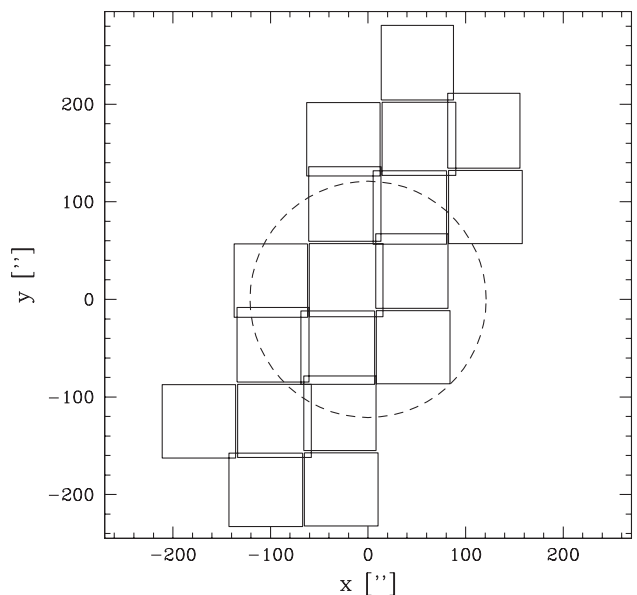


Figure 1. Layout of the observed field. The six pointings are indicated. The dashed circle indicates an aperture with a radius of $1 h_{50}^{-1} \text{ Mpc}$ centred on the Brightest Cluster Galaxy (BCG). The area covered by the observations is approximately 26.5 arcmin^2 .

been observed in two passbands. Each pointing in each filter consists of three separate short exposures, which allows an effective rejection of cosmic rays. The total integration time per pointing was 3300 s in the F606W filter, and 3200 s in the F814W filter. The reduction is described in van Dokkum et al. (2001). For the weak lensing analysis we omit the data of the Planetary Camera because the data do not reach the same depth as the Wide Field Camera. The total area covered by the observations is approximately 26.5 arcmin^2 .

Fig. 2 shows a grey-scale image of the F814W image, centred on the BCG. MS 2053 is not a rich cluster and is not easily recognized from this image. However, a colour image shows a clear enhancement of red galaxies. The image also shows the gravitationally lensed arc discovered by Luppino & Gioia (1992).

3 OBJECT ANALYSIS

The weak lensing analysis technique is based on that developed by Kaiser, Squires & Broadhurst (1995) and Luppino & Kaiser (1997), with a number of modifications that are described in detail in HFKS and HFK. We analyse each WFPC2 chip separately, and combine the object catalogues into a master catalogue once all objects have been analysed and the appropriate corrections have been applied. We use the hierarchical peak finding algorithm from Kaiser et al. (1995) to find objects with a significance $> 5\sigma$ over the local sky. These are analysed, which yields estimates for their sizes, magnitudes and shapes. As described in HFK we also estimate the error on the shape measurements, which allows a proper weighting of the sources.

The resulting catalogues are inspected visually, and spurious detections, such as diffraction spikes, HII regions in resolved galaxies, etc., are removed. We then identify the objects that are detected in both the F606W and F814W images. For these we determine colours using the same aperture for both filters. The aperture that is used scales with the Gaussian scalelength r_G of the object. This results in a sample of 2155 objects, both galaxies and

¹ Throughout this paper we will use $H_0 = 50 h_{50} \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_m = 0.3$ and $\Omega_\Lambda = 0$. This gives a scale of $1 \text{ arcsec} = 8.2 h_{50}^{-1} \text{ kpc}$ at the distance of MS 2053.

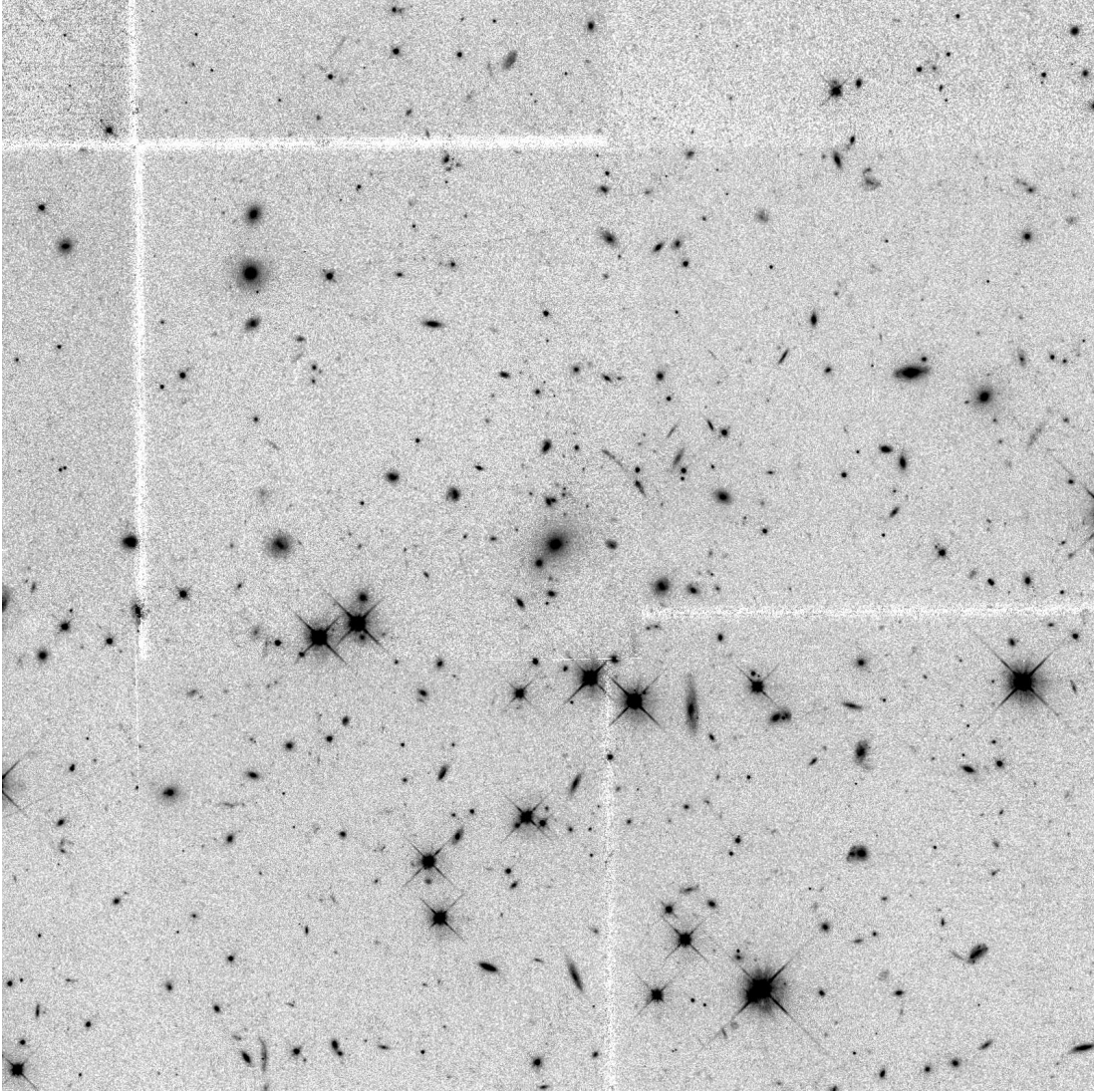


Figure 2. Grey-scale F814W image centred on the BCG. The region shown is $160 \times 160 \text{ arcsec}^2$. MS 2053 is not a rich cluster and is not easily recognized from this image. However, a colour image shows a clear enhancement of red galaxies. A gravitational lensed arc can be seen to the upper right of the cluster centre.

stars, with a corresponding number density of objects of 81 arcmin^{-2} . The objects that are detected in only one filter are small, and faint, and as a result not useful for the weak lensing analysis.

The magnitudes are zero-pointed to Vega, using the zero-points given in the *HST* Data Handbook (Voit et al. 1997). Fig. 3 shows a plot of the apparent magnitude versus the object half light radius of the detected objects in the F814W filter.

Because of the low galactic latitude of MS 2053, many stars are found in the observed field. These are located in the vertical sequence of points at a half light radius $r_h \sim 0.1 \text{ arcsec}$. The brightest stars saturate and have larger half light radii. Based on Fig. 3 we select a sample of 198 moderately bright stars. These stars are used to study the point spread function (PSF), and the results are used to correct the shapes of the faint galaxies for PSF anisotropy and the size of the PSF, as described in HFKS. The observed polarizations in the F814W images of these stars are presented in Fig. 4. HFKS studied the WFPC2 PSF using observations of the globular cluster M4, and the pattern observed here is similar to the one presented in HFKS.

The PSF changes slightly with time, and subtraction of the M4 model from the observations leaves systematic residuals. To improve the model for the PSF anisotropy, we fitted a modified model to the shape parameters of the stars in the MS 2053. It is a scaled version of the M4 model with a first-order polynomial added:

$$p_i^{\text{new}} = ap_i^{\text{M4}} + c_0 + c_1x + c_2y.$$

This model fits the observed PSF anisotropy of stars in the MS 2053 field well (the reduced χ^2 of the fit is 0.98 for 179 stars).

The next step is to determine the ‘pre-seeing’ shear polarizability P^γ (Luppino & Kaiser 1997; HFKS). The measurements of P^γ for individual galaxies are rather noisy, and therefore we bin the measurements as a function of the Gaussian scalelength r_G . Because of the poor sampling of WFPC2 images, only the shapes of galaxies with size $r_G > 0.08 \text{ arcsec}$ can be corrected reliably.²

² The peak finder program provides an estimate for r_G that is a factor $\sqrt{2}$ too large. This affects all objects, and therefore the limit listed here corresponds to the 0.12 arcsec given by HFKS.

We select objects that have $r_G > 0.08$ arcsec and remove saturated stars from the catalogues. After this selection the sample of galaxies consists of 1540 galaxies analysed from the F606W images, and 1545 from the F814W images. We note that because of this cut not all objects appear in both catalogues any longer.

Finally the shapes are corrected for the camera distortion, and the catalogues are combined into a master catalogue. We use the estimated errors on the shape measurements to combine the results

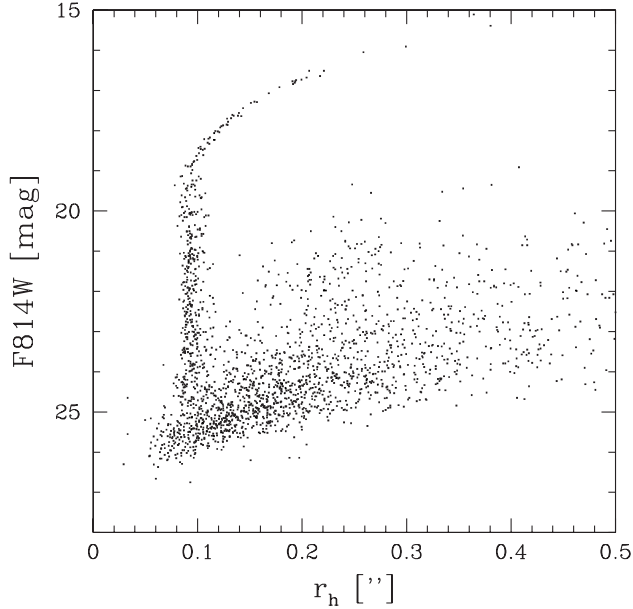


Figure 3. Plot of the apparent magnitude in the F814W filter versus half light radius r_h . Because of the low Galactic latitude of MS 2053, many stars are found, which correspond to the vertical sequence of points at $r_h \sim 0.1$ arcsec.

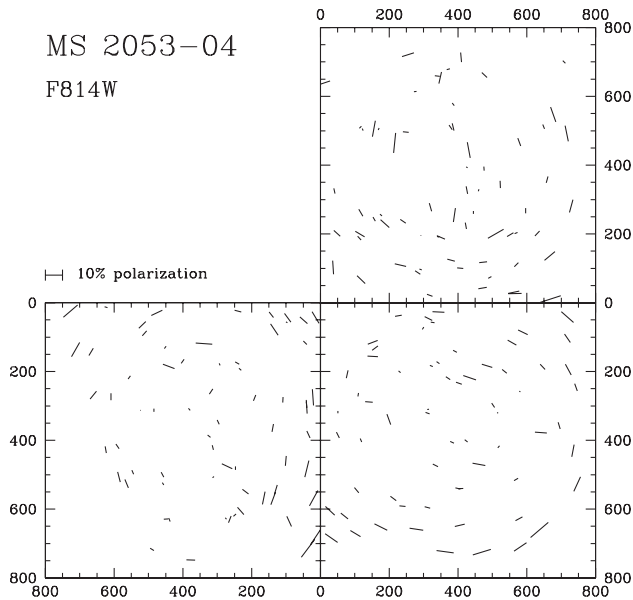


Figure 4. Observed polarization of stars selected from the F814W images. The sticks indicate the direction of the major axis of the PSF, as well as the size of the polarization. The polarizations are measured using a Gaussian weight function with a dispersion of 0.07 arcsec. The lower left panel corresponds to chip 2, the lower right to chip 3, and the upper right one denotes chip 4. We have omitted chip 1, which is the Planetary Camera.

from the F606W and F814W images in an optimal way. The resulting catalogue includes 1677 galaxies, which corresponds to a number density of galaxies of 63 arcmin^{-2} .

4 LIGHT DISTRIBUTION

Fig. 5(a) shows the colour of the galaxies in the full mosaic versus their F814W magnitude. Compared to the other two clusters for which we obtained HST mosaics, MS 2053 is less obvious from the optical images. As a result the contrast of the cluster colour–magnitude relation with the background is lower. However, in the diagram for galaxies within 1 arcmin from the BCG, the cluster colour–magnitude relation can be discerned (Fig. 5b).

To estimate the light contents of the cluster we define a sample of cluster galaxies as follows. Down to $F814W = 21$ we select spectroscopically confirmed cluster members (van Dokkum et al., in preparation). At fainter magnitudes we use the colour–magnitude relation drawn in Fig. 5(b). We select galaxies with $-0.4 < \Delta(F606W - F814W) < 0.2$ mag relative to the cluster colour–magnitude relation. To correct for contamination by field galaxies we subtract the counts from the *Hubble Deep Fields* north and south. The smoothed luminosity distribution of this sample is presented in Fig. 6(a). In Figs 6(b) and (c) grey-scale images of the smoothed number density of bright ($19.5 < F814W < 23$) and faint ($23 < F814W < 25$) galaxies are presented. Around the position of the BCG a significant overdensity of galaxies is detected. Fig. 6(a) shows most clearly that the light distribution is elongated in the direction where the arc is found (Luppino & Gioia 1992). Fig. 6(c) still shows an overdensity at the position of the cluster. Interestingly, it also shows a clear overdensity south of the cluster. The overdensity is caused by galaxies bluer than the cluster, but it is not clear whether they belong to another cluster along the line of sight.

We estimate the cluster luminosity in the rest-frame *B* band. To

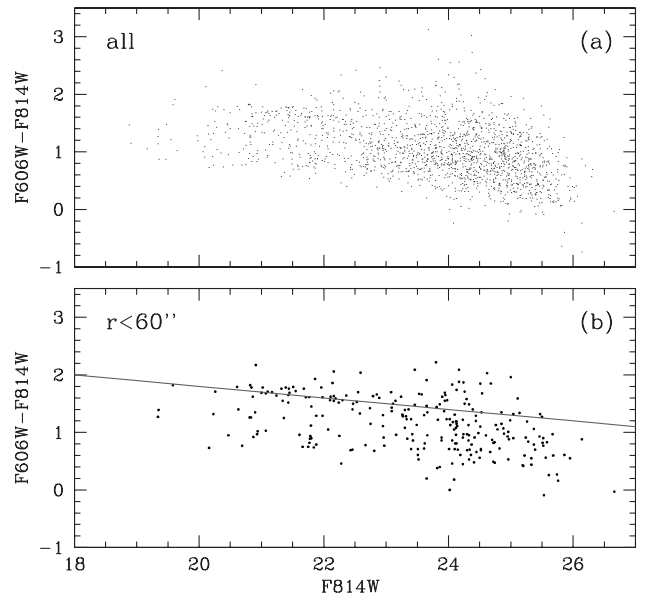


Figure 5. (a) Colour–magnitude diagram of the galaxies in the full mosaic for which colours have been determined. (b) Colour–magnitude diagram for galaxies within 1 arcmin from the BCG. The cluster colour–magnitude relation is more clearly visible in this diagram, although it is not as obvious as for other rich clusters. The line indicates the assumed cluster colour–magnitude relation.

do so we use template spectra for a range in spectral types and compute the corresponding passband correction [this procedure is similar to the method described in van Dokkum & Franx (1996)]. Thus we find the following transformation from the *HST* filters to the rest-frame *B* band:

$$B_z = F814W + 0.47(F606W - F814W) + 0.75,$$

where B_z denotes the corrected *B*-band magnitude. The luminosity is given by

$$L_B = 10^{0.4(M_{B\odot} - B_z + DM + A_{F814W})} L_{B\odot},$$

where $M_{B\odot} = 5.48$ is the solar absolute *B* magnitude, DM is the distance modulus and A_{F814W} is the extinction correction in the F814W filter towards MS 2053. The redshift of $z = 0.58$ for MS 2053 gives a distance modulus of $43.14 - 5 \log h_{50}$. We use the dust maps from Schlegel, Finkbeiner & Davis (1998) to correct for the Galactic extinction. Because of the low Galactic latitude of MS 2053, we find a rather high value of $A_{F814W} = 0.15$. We have used SEXTRACTOR (Bertin & Arnouts 1996) to determine total magnitudes for the galaxies.

The cumulative light profile as a function of distance from the cluster centre is presented in Fig. 7. The total luminosity within an aperture of radius $1 h_{50}^{-1}$ Mpc is $(3.1 \pm 0.4) \times 10^{12} h_{50}^{-2} L_{B\odot}$. The error in the luminosity reflects the uncertainty in the determination of cluster membership, and the total magnitudes measured by SEXTRACTOR. We note, however, that the error is small compared to the uncertainty in the weak lensing signal (Section 4 and further). At large radii the profile is rather steep. To compute the profile, we average the light distribution in circular bins. If the light distribution is elongated as suggested by Fig. 6(a), this leads to an overestimate of the light at large radii, where the coverage is incomplete.

5 WEAK LENSING SIGNAL

Each galaxy gives only a noisy estimate of the weak lensing signal because of its intrinsic shape. Therefore we average the shape

measurements of many sources to obtain a useful estimate of the distortion g . When we compute the ensemble-averaged distortion, we weight the contribution of each object with the inverse square of the uncertainty in the measurement of the distortion as described in HFK.

We select galaxies with $21.5 < F814W < 25.5$ and $0 < F606W - F814W < 1.4$ in our sample of background galaxies. As mentioned above, we exclude objects with sizes comparable to the PSF. The resulting sample consists of 1130 galaxies, and has a median magnitude of $F814W = 24.3$. Comparison with Fig. 5 shows that some faint cluster members might end up in this sample of background sources. The contamination will be most important in the central region of the cluster. To estimate the contamination by cluster members we examined the azimuthally averaged number density as a function of distance from the cluster centre. The profile is presented in Fig. 8. The number counts are slightly higher near the cluster centre, but the excess is not significant. In the further analysis we ignore the data inside 40 arcsec from the BCG.

The average number density of source galaxies is 43 arcmin^{-2} . Similar number densities can be reached in deep images taken from the ground (e.g. Bézecourt et al. 2000). However, the main advantage of our observations over ground-based images is the much smaller correction for the size of the PSF. As a result the lensing signal is better calibrated, and the noise in the shape measurements from *HST* data is lower.

Fig. 9(a) shows the smoothed distortion field from the sample of source galaxies (we used a Gaussian with a full width at half-maximum of $\text{FWHM} = 45 \text{ arcsec}$). The position of the BCG corresponds to the origin of the plot. A systematic tangential alignment of the sources with respect to the cluster centre can be observed.

We use the distortion field presented in Fig. 9(a) to reconstruct the projected surface mass density. The mass reconstruction has been computed using the maximum likelihood extension of the original KS algorithm (Kaiser & Squires 1993; Squires & Kaiser 1996). This algorithm has the advantage over direct inversion

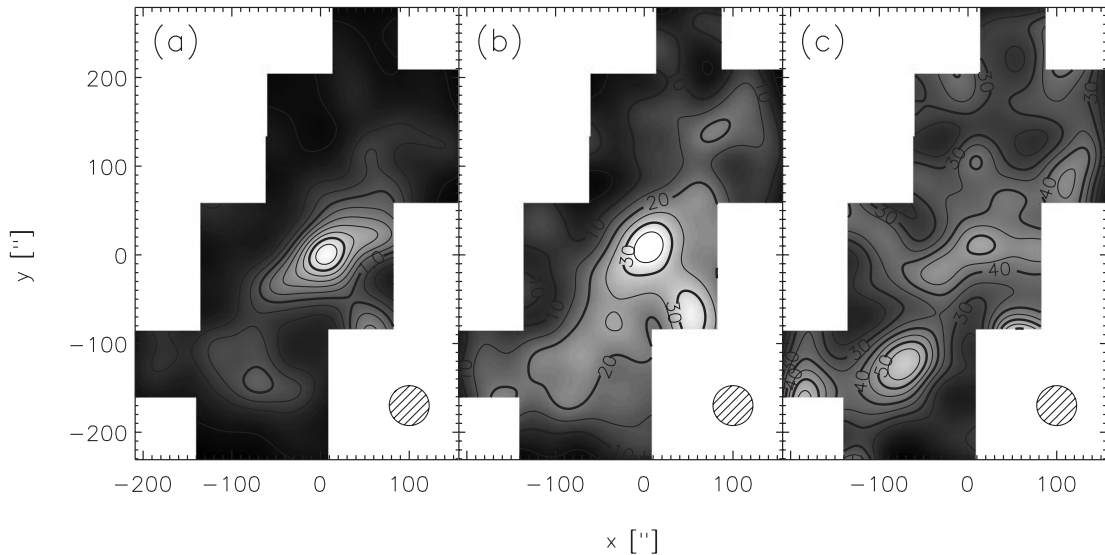


Figure 6. (a) Smoothed luminosity distribution from the sample of cluster galaxies. The definition of the sample is described in the text. The intervals between subsequent contours are $10^6 L_{B\odot} \text{ pc}^{-2}$. The position of the BCG has been used to define the cluster centre, and corresponds to the origin of this figure. The overdensity at the cluster centre is clearly visible. (b) Smoothed number density of bright galaxies ($19.5 < F814W < 23$). (c) Smoothed number density of faint galaxies ($23 < F814W < 25$). At the position of the cluster a small overdensity is still visible. Another overdensity of faint galaxies is visible to the lower left of the cluster. The number density distributions have been smoothed using a Gaussian with a FWHM of 45 arcsec (indicated by the shaded circle). In parts (b) and (c) the interval between adjacent galaxy contours is 5 arcmin^{-2} .

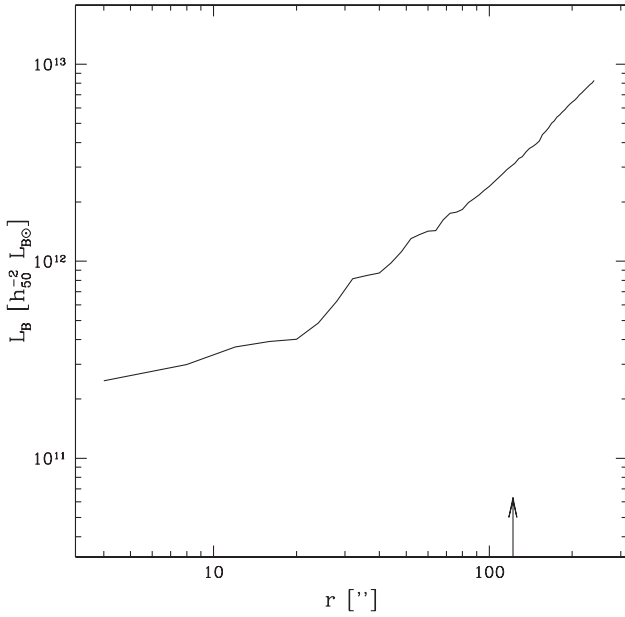


Figure 7. The cumulative rest-frame *B*-band luminosity as a function of radius from the cluster centre. The profile is calculated using the sample of cluster galaxies (selected both spectroscopically and based on colour). The arrow indicates a radius of $1 h_{50}^{-1}$ Mpc.

methods that it can be applied to fields with complicated boundaries, such as our mosaic.

Fig. 9(b) shows a grey-scale image of the reconstructed surface mass density. The peak in the mass distribution coincides with the position of the BCG. A bootstrapping resampling of the shape measurements enables us to compute the noise map of the mass reconstruction, which is presented in Fig. 9(c). The noise in the mass reconstruction increases rapidly towards the edges of the observed field. From the noise map we find that the peak in the mass distribution is detected at the 3σ level.

6 MASS AND MASS-TO-LIGHT RATIO

The azimuthally averaged tangential distortion $\langle g_T \rangle$ as a function of radius from the cluster centre is a useful measure of the lensing signal (e.g. Miralda-Escudé 1991; Tyson & Fischer 1995). The tangential distortion is defined as $g_T = -(g_1 \cos 2\phi + g_2 \sin 2\phi)$, where ϕ is the azimuthal angle with respect to the assumed cluster centre, for which we take the position of the BCG.

The azimuthally averaged tangential distortion as a function of radius from the cluster centre is presented in Fig. 10. A singular isothermal sphere model ($\kappa(r) = r_E/2r$, where r_E is the Einstein radius) gives a best-fitted $r_E = 6.3 \pm 1.8$ arcsec. To minimize the diluting effect of cluster galaxies on the weak lensing signal, we have excluded the measurements at radii smaller than 40 arcsec from the fit.

6.1 Velocity dispersion

The next step is to relate the measurement of the Einstein radius to a velocity dispersion, for which we use the photometric redshift distributions from the northern and southern *Hubble Deep Fields* (Chen et al. 1998; Fernández-Soto, Lanzetta & Yahil 1999). HFK examined the usefulness of photometric redshift distributions to calibrate the lensing signal and found that they work well.

The amplitude of the lensing signal as a function of source

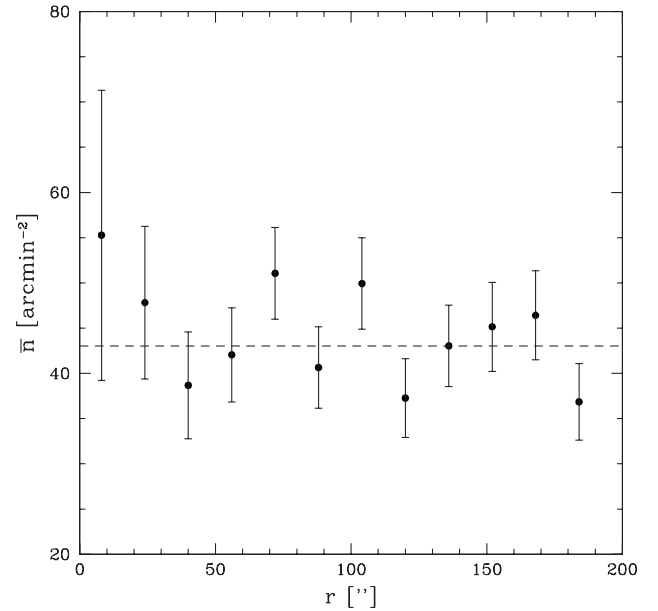


Figure 8. Azimuthally averaged number density of background galaxies ($21.5 < F814W < 25.5$ and $0 < F606W - F814W < 1.4$) as a function of distance from the cluster centre. The counts have been corrected for the boundaries of the mosaic. No significant overdensity is found at the cluster position, which indicates that the contamination by cluster members is low. The dashed line shows the average number density of background galaxies of 43 arcmin^{-2} .

redshift is characterized by β , which is defined as $\beta = \max[0, D_{ls}/D_s]$, where D_{ls} and D_s are the angular diameter distances between the lens and the source, and between the observer and the source. To compute $\langle \beta \rangle$ we also take into account that fainter galaxies are noisier and have a lower weight in the average. For our sample of sources we obtain $\langle \beta \rangle = 0.29$ (taking $\Omega_m = 0.3$ and $\Omega_\Lambda = 0$). Placing the background galaxies in a single source plane at $z = 1$ would yield a similar $\langle \beta \rangle$. Using these results we derive a velocity dispersion of $\sigma = 886_{-139}^{+121} \text{ km s}^{-1}$.

The value of $\langle \beta \rangle$ depends not only on the redshifts of the sources, but also on the cosmological parameters that define the angular diameter distances. For an $\Omega_m = 1$ and $\Omega_\Lambda = 0$ model we find essentially the same $\langle \beta \rangle$ and $\sigma = 881_{-138}^{+120} \text{ km s}^{-1}$. In a Ω_Λ -dominated universe the changes are larger. Assuming $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$ gives $\langle \beta \rangle = 0.33$, and results in $\sigma = 831_{-130}^{+113} \text{ km s}^{-1}$.

The result from the weak lensing analysis is in excellent agreement with the observed velocity dispersion of $817 \pm 80 \text{ km s}^{-1}$, which was determined from the velocities of 52 cluster members (van Dokkum et al., in preparation).

Clowe (1998) obtained deep *R*-band images of MS 2053 with the Keck Telescope, and measured the weak lensing signal. He used a source redshift of $z = 1.75$ and derived a velocity dispersion of $\sigma \sim 700 \text{ km s}^{-1}$. Such high average source redshifts are unrealistic, in particular when compared to the $z = 1$ we use, based on photometric redshift distributions. For more realistic redshift distributions, the result of Clowe (1998) increases to $\sigma \sim 900 \text{ km s}^{-1}$, in good agreement with our results.

Luppino & Gioia (1992) discovered a gravitationally lensed blue arc in deep images of MS 2053. The arc is located approximately 15 arcsec north of the BCG. To date, the redshift of the arc is not known. If we assume a redshift of $z = 2$ for the arc, and adopt an SIS model (where the position of the arc gives the Einstein radius), the corresponding velocity dispersion is about 1030 km s^{-1} . This

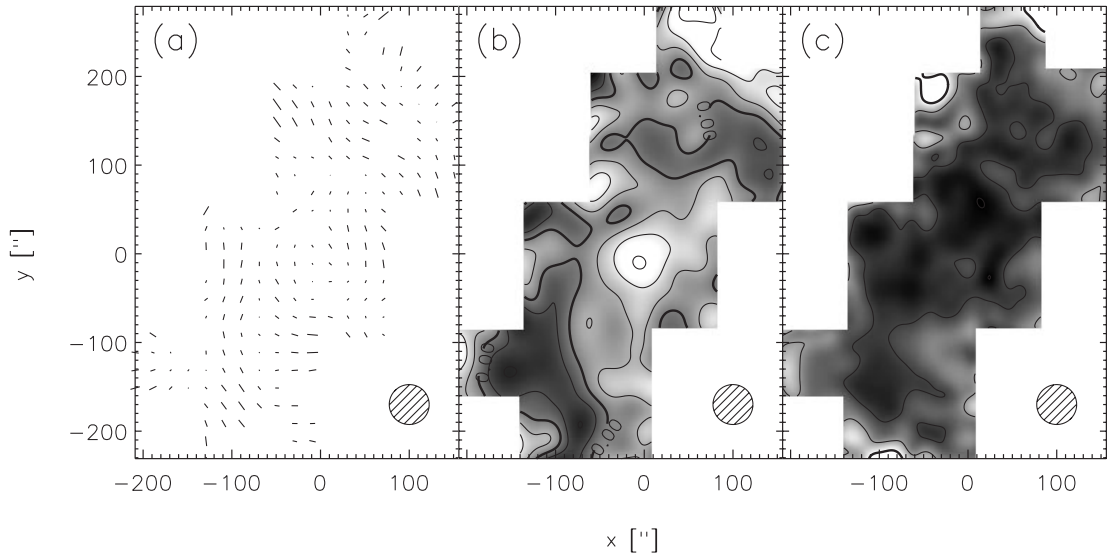


Figure 9. (a) Smoothed distortion field g obtained using galaxies with $22 < F814W < 26$ and $0 < F606W - F814W < 1.6$. The measurements have been smoothed using a Gaussian with a FWHM of 45 arcsec (indicated by the shaded circle). The orientation of the sticks indicates the direction of the distortion, and the length is proportional to the amplitude of the signal. The origin of the plot coincides with the assumed cluster centre. (b) The corresponding reconstruction of the projected surface mass density. The interval between adjacent contours is 0.05 in κ . (c) Noise map of the mass reconstruction from bootstrapping resampling of the shape measurements. It shows that the noise in the reconstruction increases rapidly towards the edges of the observed field. The interval between adjacent contours is 0.025. The peak in the mass reconstruction is detected at the 3σ level.

value is higher than, but consistent with, the weak lensing estimate. Moreover, if the mass distribution is elongated in the direction of the arc, the strong lensing mass estimate is lowered. Fig. 6(a) indicates that the light distribution is elongated, roughly in the direction of the giant arc. Because of the low signal-to-noise ratio of the weak lensing signal, we cannot constrain the elongation of the mass distribution.

6.2 Mass-to-light ratio

From the sample of cluster galaxies we estimate a total cluster luminosity of $(3.1 \pm 0.4) \times 10^{12} h_{50}^{-2} L_{B\odot}$ within an aperture of radius $1 h_{50}^{-1}$ Mpc (see Section 4). The best-fitting SIS model gives a projected mass of $(5.7 \pm 1.6) \times 10^{14} h_{50}^{-1} M_{\odot}$ in the same aperture. Thus we obtain an average mass-to-light ratio of $184 \pm 56 h_{50} M_{\odot}/L_{B\odot}$ within $1 h_{50}^{-1}$ Mpc.

Kelson et al. (1997) have studied the fundamental plane of MS 2053. They find that the early-type galaxies in the cluster define a clear fundamental plane. Comparison with low-redshift clusters suggests that the structure of early-type galaxies has changed little since $z = 0.58$. Similar analyses have been performed for other clusters (e.g. van Dokkum & Franx 1996; van Dokkum et al. 1998) and have shown that the mass-to-light ratios of early-type galaxies evolve with redshift, which is attributed to luminosity evolution.

As a result also the global cluster mass-to-light ratios evolve with redshift. The mass-to-light ratio of early-type galaxies in MS 2053 in the B band is 37 ± 4 per cent lower than present-day values (Kelson et al. 1997). Under the assumption that the total luminosity of the cluster has changed by the same amount, we find an average mass-to-light ratio within $1 h_{50}^{-1}$ Mpc of $291 \pm 89 \pm 19 h_{50} M_{\odot}/L_{B\odot}$, corrected for luminosity evolution to $z = 0$. The first contribution to the error budget is the statistical uncertainty in the determination of the mass-to-light ratio, and the second contribution is due to the uncertainty in the correction for luminosity evolution.

Under the assumption that the light traces the mass, we can

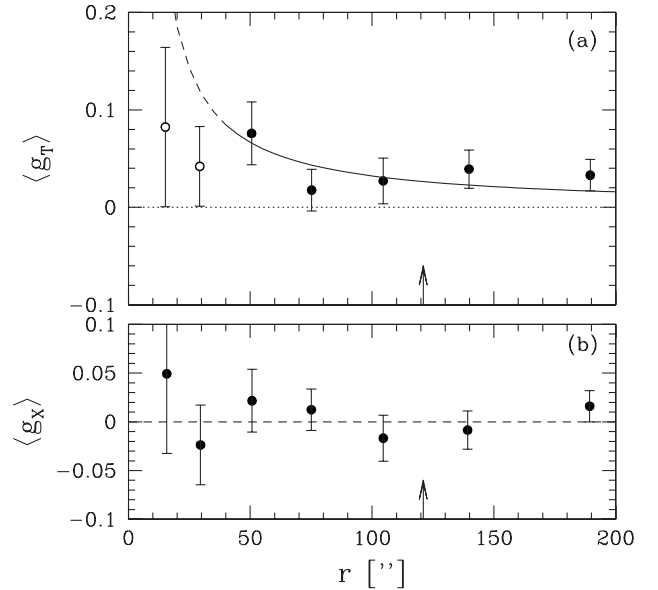


Figure 10. (a) Average tangential distortion as a function of radius from the cluster centre, for which we took the position of the BCG. The line corresponds to the profile of a singular isothermal model ($\kappa(r) = r_E/2r$, where r_E is the Einstein radius) fitted to the data. Because of possible contamination by faint cluster members we only fit to the data at radii larger than 40 arcsec. The best-fitting value for the Einstein radius is $r_E = 6.2 \pm 1.8$ arcsec. (b) Average signal when the phase of the distortion is increased by $\pi/2$. If the signal shown in (a) is caused by gravitational lensing, $\langle g_x \rangle$ should vanish, as is observed. In both figures, the arrows indicate a radius of $1 h_{50}^{-1}$ Mpc.

derive the expected tangential distortion as a function of radius. To measure the mass-to-light ratio, we scale the computed tangential distortion g_T^{lum} to match the observed signal. In Fig. 11(a) the resulting profile (solid line) is shown. The ratio of the computed and observed signal is presented in Fig. 11(b). Because of possible

contamination by faint cluster members we exclude the points at radii less than 40 arcsec from the fit. We find that the results are consistent with a constant mass-to-light ratio with radius, and we find an average value of $195 \pm 58 h_{50} M_{\odot}/L_{B\odot}$.

7 COMBINED RESULTS FROM RICH CLUSTERS

With the analysis of MS 2053 we have a sample of three clusters for which the mass distribution has been studied using mosaics of WFPC2 images. For the analysis in Section 7.3, we augment this sample with the $z = 0.22$ cluster Abell 2219, which has been studied from the ground by Bézeacourt et al. (2000).

The mass distribution of low-redshift clusters ($z < 0.4$) can be measured well from the ground: most of the lensing signal comes from source galaxies with $z > 0.5$, for which accurate shape measurements can be made. Only for high-redshift clusters ($z > 0.5$) are *HST* observations important, because in these cases the relevant source galaxies have redshifts $z > 1$.

In addition, A2219 has been analysed in the same way as the three clusters imaged by *HST*. The shape measurements have been corrected for PSF anisotropy and the circularization by the PSF. Furthermore, we have used photometric redshifts to calibrate the mass of the cluster, and determined the *B*-band luminosity in the same way as was done for the other clusters. As a result the A2219 results can be compared directly to the other three clusters in the sample.

All clusters were included on the basis of their X-ray properties, which are listed in Table 1. We note that by selecting clusters based on their X-ray properties (i.e. selecting X-ray luminous clusters), we might be biased towards more relaxed clusters, which in turn might result in tight relations between the cluster properties, such as mass, galaxy velocity dispersion, X-ray temperature and luminosity. Detailed studies are needed to investigate whether the results obtained for X-ray luminous clusters hold for the more general population of clusters of galaxies.

7.1 Comparison between weak lensing mass and dynamical mass

The large number of spectroscopic confirmed members in each of the remaining three clusters results in accurate measurements of their galaxy velocity dispersions. In this section we compare the weak lensing estimates of the cluster velocity dispersions to the velocity dispersion of the galaxies. In Table 2 we list the results of the best-fitting SIS model to the observed weak lensing signal, as well as the corresponding velocity dispersion inferred from lensing.

For all three clusters we use the photometric redshift distribution from the *Hubble Deep Fields* (Chen et al. 1998; Fernández-Soto et al. 1999). This lowered the $\langle \beta \rangle$ for Cl 1358 slightly compared to the value used in HFKS, and the new estimate for the weak lensing velocity dispersion is listed in Table 2.

The comparison with the velocity dispersions of cluster galaxies shows a good agreement in all three cases, suggesting that the galaxy velocity dispersions are characteristic of the cluster as a whole. Similar comparisons have been made in the past (e.g. Smail et al. 1997; Wu et al. 1998; Allen 1998). The samples of clusters used by Wu et al. (1998) and Allen (1998) are rather inhomogeneous: different methods were used for the correction of the circularization of the background sources, and no realistic redshift distributions of the sources have been used. Wu et al. (1998) and

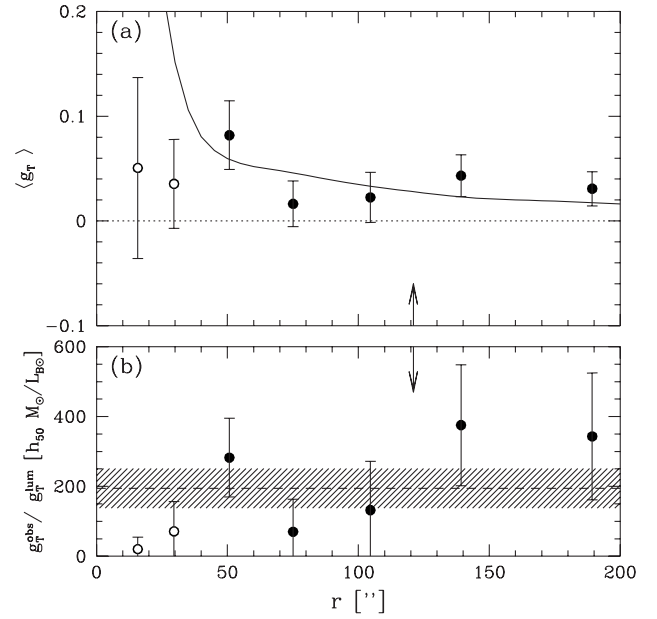


Figure 11. (a) Plot of the average tangential distortion as a function of radius from the BCG. The solid line is the expected tangential distortion (scaled by the mass-to-light ratio to fit the observations) derived from the average radial light profile, assuming the mass-to-light ratio is constant with radius. (b) The ratio of the observed distortion and the derived distortion from the light (taking $M/L_B = 1$ in solar units). The shaded region indicates the 1σ region around the average of the points with radii larger than 40 arcsec. The observations are consistent with a constant mass-to-light ratio of $196 \pm 58 h_{50} M_{\odot}/L_{B\odot}$. In both figures, the arrows indicate a radius of $1 h_{50}^{-1}$ Mpc.

Table 1. X-ray properties of the three clusters for which we have obtained *HST* mosaics, as well as the properties of A2219.

| | z | L_X (2–10 keV) ($h_{50}^{-2} 10^{44} \text{ erg s}^{-1}$) | kT (keV) | Ref. ^a |
|------------|------|--|----------------------|-------------------|
| A2219 | 0.22 | 38 | 9.5 ± 0.6 | 1 |
| Cl 1358+62 | 0.33 | 11.4 ± 0.3 | 6.9 ± 0.5 | 2 |
| MS 2053–04 | 0.58 | 7.9 ± 0.7 | $8.1^{+3.7}_{-2.2}$ | 2 |
| MS 1054–03 | 0.83 | 28.6 | $12.4^{+3.1}_{-2.2}$ | 3 |

^aReferences: (1) Allen (1998); (2) Henry (2000); (3) Donahue et al. (1998).

Allen (1998) find in general a fair agreement between the velocity dispersions of the galaxies and the velocity dispersions derived from weak lensing analyses.

A more systematic study was presented by Smail et al. (1997), who determined the weak lensing signal of 12 distant clusters observed with WFPC2. Each cluster was observed with one pointing on the cluster core. Smail et al. (1997) computed the average tangential distortion within an annulus $120 < r < 400 h_{50}^{-1}$ kpc, and plotted the result against the observed cluster dispersion. Under the assumption that the cluster mass distribution is described by an SIS model, this measurement provides an estimate for the velocity dispersion. Fig. 12 shows the weak lensing estimate of the velocity dispersion versus the velocity dispersion of cluster members for our sample. The large open points are the results from our *HST* mosaics, when we confine the weak lensing analysis to the annulus used by Smail et al. (1997). The large solid points show the results when the full *HST* mosaic is used for the

Table 2. Results from the weak lensing analyses of the three clusters for which *HST* mosaics were obtained.

| | z | r_E (arcsec) | $\langle\beta\rangle$ | σ (WL) (km s ⁻¹) | σ (galaxies) (km s ⁻¹) | Ref. ^a |
|------------|------|-------------------|-----------------------|--|--|-------------------|
| A2219 | 0.22 | 17.4 ± 2.0 | 0.46 | 1075^{+61}_{-63} | — | — |
| Cl 1358+62 | 0.33 | 10.8 ± 1.4 | 0.56 | 835^{+52}_{-56} | 910 ± 54 | 1 |
| MS 2053-04 | 0.58 | 6.3 ± 1.8 | 0.29 | 886^{+121}_{-139} | 817 ± 80 | 2 |
| MS 1054-03 | 0.83 | 11.5 ± 1.4 | 0.23 | 1311^{+83}_{-89} | 1150 ± 90 | 3 |

^aReferences: (1) Carlberg et al. (1997); (2) van Dokkum et al. (in preparation); (3) van Dokkum (1999).

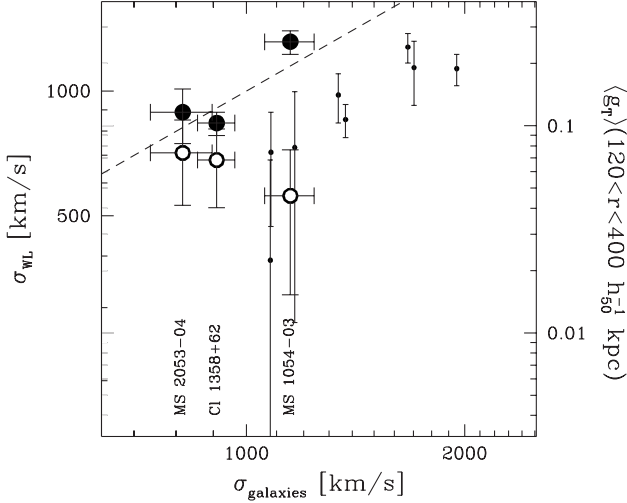


Figure 12. Weak lensing velocity dispersion (as inferred from the best-fitting SIS model) versus the observed velocity dispersion of the galaxies. The large open points are the results from our *HST* mosaics, measuring the distortion in an annulus $120 < r < 400 h_{50}^{-1}$ kpc. The results from Smail et al. (1997) are measured in the same annulus and are indicated by the small dots. The large solid points correspond to the measurements when the complete *HST* mosaics are used. The line of equality between the weak lensing velocity dispersion and the velocity dispersion of cluster galaxies is indicated by the dashed line. Our large-field weak lensing results agree well with this line. To allow a direct comparison with fig. 4 from Smail et al. (1997) we have indicated the average tangential distortion in an annulus $120 < r < 400 h_{50}^{-1}$ on the right-hand axis.

weak lensing analysis. The right-hand vertical axis displays the value of the average tangential distortion in the annulus used by Smail et al. (1997), and allows a direct comparison with their fig. 4. For comparison, we also show the results from Smail et al. (1997) (small dots).

Smail et al. (1997) found a discrepancy between their weak lensing signal and the velocity dispersion of the galaxies. They argued that the velocity dispersions from the galaxies are overestimated by ~ 40 per cent, compared to the velocity dispersion expected from the weak lensing analysis. To derive the cluster velocity dispersions, Smail et al. (1997) used the spectroscopic samples from Dressler & Gunn (1992), which were aimed at identifying blue cluster members. Smail et al. (1997) found that the typical cluster dispersion from early-type galaxies is smaller compared to the whole sample of galaxies.

When we confine the weak lensing analysis of our *HST* mosaics to the same annulus used by Smail et al. (1997), we too find that the velocity dispersion inferred from the weak lensing analysis is lower than the velocity dispersion of the galaxies. However, the results

based on the full mosaics agree well with the line of equality (dashed line).

This indicates that limiting the weak lensing analysis to the cluster core results in a systematic underestimate of the cluster velocity dispersion. The largest change is seen for MS 1054. For this cluster the explanation is straightforward. HFK showed that the mass distribution in the cluster centre is complex, consisting of three distinct clumps. As a result the average tangential distortion is lowered, when only the inner $400 h_{50}^{-1}$ kpc are considered in the analysis.

Several effects can introduce a systematic offset between the weak lensing results and the dynamical measurements (e.g. Smail et al. 1997). Because we find a good agreement between the two estimates when the lensing signal is measured from wide-field data, we argue that the cluster mass profile in the core gives rise to a discrepancy between the velocity dispersion of galaxies and the weak lensing measurements. Only if the profile is isothermal does one expect a good agreement, but substructure or a shallower density profile lowers the lensing signal compared to the value expected from the SIS model. However, we note that the selection of galaxies used to derive the galaxy velocity dispersion in Smail et al. (1997) also contributes to the reported discrepancy.

7.2 Average cluster mass profile

Numerical simulations have indicated that dark matter haloes originating from dissipationless collapse of density fluctuations may follow a universal density profile (e.g. Navarro, Frenk & White 1997) (NFW). The NFW profile appears to be an excellent description of the radial mass distribution in these simulations. The NFW profile is given by

$$\rho(r) = \frac{\delta_c \rho_c}{(r/r_s)(1 + r/r_s)^2}, \quad (1)$$

where ρ_c is the critical density of the universe, δ_c is the characteristic overdensity and r_s is the scale radius given by $r_s = r_{200}/c$, which all depend on the redshift and mass of the halo. The parameter c is referred to as the concentration parameter. Given the cosmology, redshift and mass of the halo, r_{200} follows immediately, and the values of δ_c and c can be computed using the routine CHARDEN made available by Julio Navarro.³

We have fitted the predicted profiles from NFW haloes to the observed tangential distortion of each cluster, and the best-fitting parameters are listed in Table 3. We have used a value of $\Gamma = 0.18$ for the shape parameter of the cold dark matter (CDM) power spectrum. The only parameter that we fitted is M_{200} , the mass

³The routine CHARDEN can be obtained from <http://pinot.phys.uvic.ca/~jfn/charden>

Table 3. Best-fitting parameters for the NFW profile (columns 3–6). The fitted parameter is M_{200} and the other parameters are computed using CHARDEN, for an Λ CDM cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0$ and shape parameter $\Gamma = 0.18$. The errors indicate the 68 per cent confidence intervals. Columns 7 and 9 give the χ^2 for the best-fitting NFW profile and SIS model, respectively. The probabilities of finding a larger value for χ^2 are listed in columns 8 and 10.

| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|------------|------|--|---|------------------------|---------------------------------------|-----------------------|-----------------------------------|-----------------------|-----------------------------------|
| | z | M_{200} ($10^{14} h_{50}^{-1} M_\odot$) | r_{200} ($h_{50}^{-1} \text{Mpc}$) | c | r_s ($h_{50}^{-1} \text{kpc}$) | χ^2_{NFW} | $P(\chi^2 > \chi^2_{\text{NFW}})$ | χ^2_{SIS} | $P(\chi^2 > \chi^2_{\text{SIS}})$ |
| A2219 | 0.22 | $16.6^{+3.4}_{-3.2}$ | $2.63^{+0.17}_{-0.18}$ | $5.01^{+0.13}_{-0.11}$ | 524^{+47}_{-48} | 8.6 | 0.74 | 10.0 | 0.62 |
| Cl 1358+62 | 0.33 | $2.8^{+0.8}_{-0.8}$ | $1.36^{+0.12}_{-0.14}$ | $5.86^{+0.19}_{-0.14}$ | 231^{+27}_{-31} | 15.1 | 0.24 | 15.7 | 0.20 |
| MS 2053–04 | 0.58 | $2.8^{+1.6}_{-1.5}$ | $1.18^{+0.19}_{-0.27}$ | $5.46^{+0.40}_{-0.22}$ | 217^{+46}_{-61} | 8.6 | 0.74 | 9.6 | 0.65 |
| MS 1054–03 | 0.83 | $14.4^{+2.8}_{-2.6}$ | $1.81^{+0.11}_{-0.12}$ | $4.46^{+0.09}_{-0.08}$ | 407^{+32}_{-34} | 11.3 | 0.13 | 11.5 | 0.12 |

enclosed within a sphere of radius r_{200} , and the other parameters are the ones produced by CHARDEN given M_{200} . The errors for r_{200} , c and r_s listed in Table 3 only reflect the uncertainty in these parameters because of the uncertainty in the measurement of M_{200} . We note that the resulting parameters are mainly determined by the amplitude of the lensing signal (i.e. the mass of the halo) and not by the shape of the density profile. Because of the strong substructure in the centre of MS 1054 we excluded the measurements at radii less than 75 arcsec.

We now examine whether the NFW predictions match the actual observations. To do so, we scale the amplitude of the tangential distortion profiles of the four clusters to the signal of a cluster with M_{200} of $5 \times 10^{14} h_{50}^{-1} M_\odot$ at a redshift $z = 0.5$ (where we placed the sources at infinite redshift), and scale the data radially in units of the derived value of r_s (listed in Table 3), the scalelength of the NFW profile. The resulting ensemble-averaged tangential distortion as a function of r_s is presented in Fig. 13(a). This figure also shows the best-fitting SIS model (dashed line) and the best-fitting NFW profile (solid line).

The NFW profile provides a good fit to the data [$\chi^2 = 12.8$; $P(\chi^2 > 12.8) = 0.38$]. The SIS model fit is worse with a $\chi^2 = 17.1$ [$P(\chi^2 > 17.1) = 0.15$]. The NFW model that is fitted to the observations has a concentration parameter that is $\alpha \times c_{\text{OCDM}}$. Thus we test whether the predicted concentration parameters agree with the observed lensing signal. Fig. 13(b) shows $\Delta\chi^2 = \chi^2 - \chi^2_{\text{min}}$ as a function of α in units of c_{OCDM} . We find that the best-fitting value is $0.79^{+0.44}_{-0.15} \times c_{\text{OCDM}}$. Thus the predicted concentration parameter is in good agreement with the observations.

In the above, we have used NFW models for which the parameters were obtained from the lensing data. Although the parameters are essentially determined by the amplitude of the lensing signal, it is useful to examine this in more detail. To do so, we use the observed velocity dispersions of the galaxies to obtain an estimate of r_{200} , using $r_{200} = \sqrt{3} \sigma_{\text{gal}} / 10H(z)$, where $H(z)$ is the value of the Hubble parameter at the redshift of the cluster. We omit A2219, because the galaxy velocity dispersion is not known. We compute the NFW parameters and scale the tangential distortion profiles of the three remaining clusters. The resulting ensemble-averaged profile is compared to the NFW profile in the same way as before. We find that the NFW profile is a good fit ($\chi^2 = 13.4$), and that the SIS model fits worse ($\chi^2 = 20.5$). The best-fitting concentration parameter is found to be $0.83^{+0.47}_{-0.37} \times c_{\text{OCDM}}$. Thus both approaches yield similar results.

These results show that a systematic weak lensing study of a number of clusters provides a direct way to test consistency of the predictions of the theory of dissipationless collapse in Λ CDM cosmologies.

7.3 Cluster mass-to-light ratio

Table 4 lists the estimates of the average mass-to-light ratio within an aperture of $1 h_{50}^{-1} \text{Mpc}$ radius. All values are given in the rest-frame B band. Except for A2219, the luminosity evolution of the early-type galaxies has been measured by studying the fundamental plane (e.g. Kelson et al. 1997; van Dokkum et al. 1998). Under the assumption that the global cluster mass-to-light ratio evolves similarly with redshift, we can correct the observed mass-to-light ratio for luminosity evolution to $z = 0$, and the results are also listed in Table 4. The uncertainty in the luminosity evolution results in an additional contribution to the total error budget. In Table 4 we list the statistical error in the measurement of the mass-to-light ratio and the uncertainty due to the correction for luminosity evolution separately.

Fig. 14 shows the observed average mass-to-light ratio in the B band within an aperture of $1 h_{50}^{-1} \text{Mpc}$ as a function of cluster redshift. The shaded region in Fig. 14 corresponds to the average of the observed mass-to-light ratios (i.e. assuming no luminosity evolution), which yields a value of $M/L_B = 151 \pm 12 h_{50} M_\odot/L_{B\odot}$. The observations are inconsistent with an unevolving cluster mass-to-light ratio that does not evolve at the 99 per cent level.

The solid line corresponds to the luminosity evolution as a function of redshift as inferred from studies of the fundamental plane of distant clusters of galaxies (van Dokkum et al. 1998), scaled to fit the observed total cluster mass-to-light ratios. The evolution of the cluster mass-to-light ratio of X-ray selected clusters is consistent with the evolution of the mass-to-light ratio of the early-type galaxies. Van Dokkum et al. (1998) found that the M/L_B ratio evolves as $\Delta \log M/L_B \propto (-0.40 \pm 0.04)z$, which results in an average value of $M/L_B = 239 \pm 18 \pm 9 h_{50} M_\odot/L_{B\odot}$ for clusters at $z = 0$. The first error indicates the statistical uncertainty in the measurement of the mass-to-light ratio, and the second error indicates the additional error introduced by the uncertainty in the luminosity evolution.

Carlberg, Yee & Ellingson (1997) analysed a sample of 16 rich clusters, and also found that the cluster mass-to-light ratios are consistent with a universal value. They found an average value of $M/L_r = 119 \pm 21 h_{50} M_\odot/L_{r\odot}$. To convert this value to a mass-to-light ratio in the B band, we assume an average colour of the cluster of $B - r = 1.07$, which corresponds to the typical colour of S0 galaxies (Jørgensen, Franx & Kjaergaard 1995). Thus we find that the estimate for the average cluster mass-to-light ratio from Carlberg et al. (1997) corresponds to $219 \pm 38 h_{50} M_\odot/L_{B\odot}$ (where we also corrected for luminosity evolution to $z = 0$), in excellent agreement with our results.

Given the small spread in cluster mass-to-light ratios the star

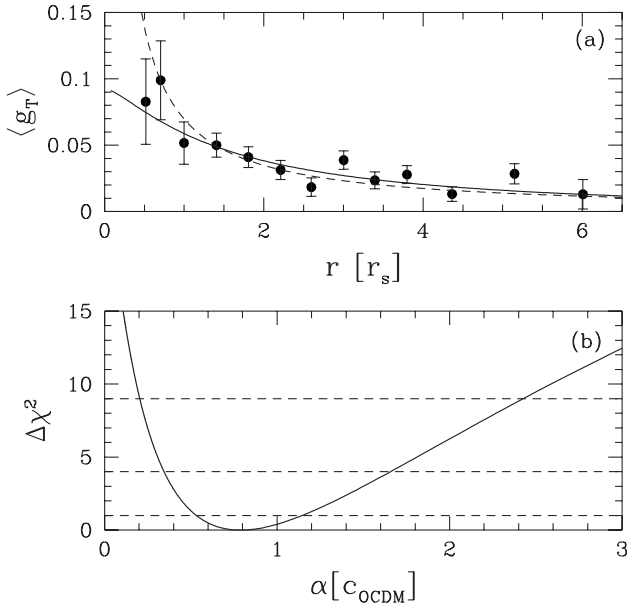


Figure 13. (a) The ensemble-averaged tangential distortion as a function of radius (in units of $r_s = r_{200}/c$) for the four clusters in the sample. We used the concentration parameters listed in Table 3. The solid line is the best-fitting NFW profile, and the dashed line is the best-fitting SIS model. The NFW provides the best-fitting ($\chi^2 = 12.8$), whereas the SIS model fit is worse ($\chi^2 = 17.1$). (b) $\Delta\chi^2 = \chi^2 - \chi^2_{\min}$ as a function of ratio between the measured and predicted concentration parameter α/c_{OCDM} . The best-fitting NFW profile yields an observed concentration parameter that is $0.79^{+0.44}_{-0.15} \times c_{\text{OCDM}}$.

formation efficiency in rich clusters appears to be a well-regulated process, although the sample of clusters needs to be increased before firmer conclusions can be drawn.

8 CONCLUSIONS

We have presented the results of our weak lensing analysis of MS 2053–04, a cluster of galaxies at a redshift $z = 0.58$, for which we detect a clear lensing signal. It is the third cluster we have studied using a two-colour mosaic of deep WFPC2 images. Previously we have studied Cl 1358+62 ($z = 0.33$; HFKS) and MS 1054–03 ($z = 0.83$; HFK).

The selected sample of background sources (with $21.5 < F814W < 25.5$ and $0 < F606W - F814W < 1.4$) has a number density of galaxies of 43 arcmin^{-2} . Similar number densities can be reached in deep ground-based observations, but the correction for the circularization by the PSF in WFPC2 images is much smaller. As a result the lensing signal can be measured more accurately from space-based images.

The position of the peak in the reconstruction of the cluster mass surface density agrees well with the peak in the light distribution. To measure the mass of the cluster we fit an SIS model to the observed azimuthally averaged tangential distortion. The corresponding value for the Einstein radius is $r_E = 6.2 \pm 1.8 \text{ arcsec}$. To relate the Einstein radius to an estimate of the cluster velocity dispersion, we use published photometric redshift distributions inferred from the northern and southern *Hubble Deep Fields*. The best-fitting SIS model corresponds to a velocity dispersion of $\sigma = 886^{+121}_{-139} \text{ km s}^{-1}$, which is in excellent agreement with the observed velocity dispersion of cluster galaxies of $817 \pm 80 \text{ km s}^{-1}$.

We have analysed the weak lensing signal of three clusters using wide-field *HST* data, and we find that the velocity dispersion

Table 4. Average mass-to-light ratio within apertures of $1 h_{50}^{-1} \text{ Mpc}$ radius. The error budget of the mass-to-light ratios corrected for luminosity evolution consists of the statistical uncertainty in the measurement of the mass-to-light ratio (first error) and the contribution due to the uncertainty in the luminosity evolution (second error).

| | z | M/L_B (obs) ($h_{50} \text{ M}_\odot/L_{B\odot}$) | M/L_B ($z = 0$) ($h_{50} \text{ M}_\odot/L_{B\odot}$) |
|------------|------|--|--|
| A2219 | 0.22 | 210 ± 24 | $256 \pm 29 \pm 6$ |
| Cl 1358+62 | 0.33 | 141 ± 23 | $186 \pm 30 \pm 10$ |
| MS 2053–04 | 0.58 | 184 ± 56 | $291 \pm 89 \pm 19$ |
| MS 1054–03 | 0.83 | 124 ± 17 | $269 \pm 37 \pm 31$ |

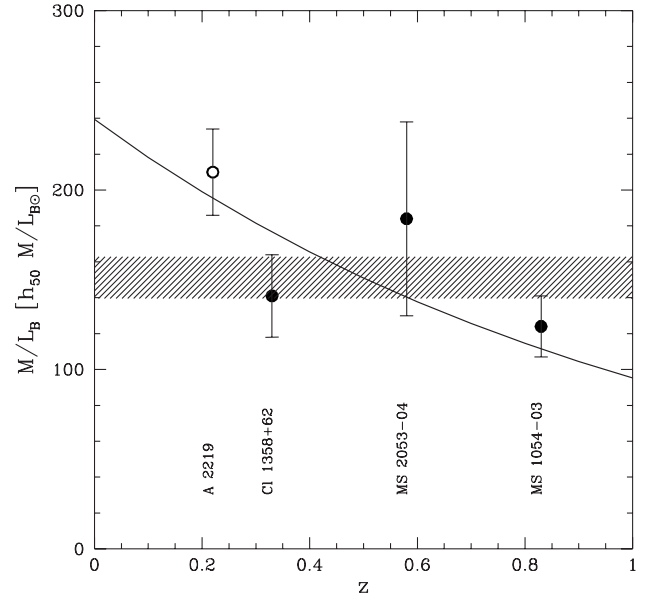


Figure 14. The observed average mass-to-light ratio within a $1 h_{50}^{-1} \text{ Mpc}$ radius aperture of the clusters in the sample as a function of redshift. The shaded region indicates the 1σ region around the average mass-to-light ratio (assuming no luminosity evolution). The assumption of an unevolving mass-to-light ratio is excluded at the 99 per cent confidence level. The solid line corresponds to the luminosity evolution as a function of redshift as inferred from studies of the fundamental plane of distant clusters of galaxies (van Dokkum et al. 1998), and is consistent with the observations.

derived from weak lensing agrees well with the velocity dispersion of the cluster galaxies. This result differs from those of Smail et al. (1997) who compared the weak lensing signal to the galaxy velocity dispersion using *HST* observations of cluster cores. Based on our results we argue that the discrepancy is caused by deviations from the SIS model in the inner regions of clusters (substructure or a flatter profile). To obtain an accurate estimate of the weak lensing velocity dispersion wide-field data are necessary.

We study the average cluster profile derived from a sample of four clusters that have been studied through weak lensing. The NFW profile fits the ensemble-averaged lensing signal well, and the predicted concentration parameter is in good agreement with the observations: the observed value is found to be $0.79^{+0.44}_{-0.15}$ times the predicted value for an OCDM model.

The observed average mass-to-light ratio of MS 2053 within a $1 h_{50}^{-1} \text{ Mpc}$ radius aperture is $184 \pm 56 h_{50} \text{ M}_\odot/L_{B\odot}$. We have examined the mass-to-light ratios of the clusters in our sample, and find that the results are inconsistent with a non-evolving universal mass-to-light ratio. The measurements are consistent with a

universal mass-to-light ratio for rich, X-ray selected, clusters of galaxies which evolves with redshift similarly to the luminosity evolution of the cluster galaxies (e.g. Kelson et al. 1997; van Dokkum et al. 1998). The average cluster mass-to-light ratio, corrected to $z=0$, is found to be $M/L_B = 239 \pm 18 \pm 9 h_{50} M_\odot/L_{B\odot}$ (where the first error indicates the statistical uncertainty in the measurement of the mass-to-light ratio, and the second error is due to the uncertainty in luminosity evolution), in good agreement with the results from Carlberg et al. (1997) based on a dynamical study of 16 rich clusters. The small spread in cluster mass-to-light ratios suggests that the total star formation in clusters is a well-regulated process.

ACKNOWLEDGMENTS

The work reported herein is based on observations with the NASA/ESA *Hubble Space Telescope* obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. PGvD was supported by Hubble Fellowship grant HF-01126.01-99A. We would like to thank the Kapteyn Astronomical Institute for their generous support.

REFERENCES

- Allen S. W., 1998, MNRAS, 296, 392
 Bahcall N. A., Fan X., 1998, ApJ, 504, 1
 Bertin E., Arnouts S., 1996, A&AS, 117, 393
 Bézecourt J., Hoekstra H., Gray M. E., Abseil Salam H. M., Kuijken K., Ellis R. S., 2000, A&A, submitted, astro-ph/0001513
 Carlberg R. G., Yee H. C., Ellingson E., 1997, ApJ, 478, 462
 Chen H.-W., Fernández-Soto A., Lanzetta K. M., Pascarelle S. M., Puetter R. C., Yahata N., Yahil A., 1998, preprint, astro-ph/9812339
 Clowe D. L., 1998, PhD thesis, Univ. Hawaii
 Della Ceca R., Scaramella R., Gioia I. M., Rosati P., Fiore F., Squires G., 2000, A&A, 353, 498
 Donahue M., Voit G. M., Gioia I., Luppino G. A., Hughes J. P., Stocke J. T., 1998, ApJ, 502, 550
 Dressler A., Gunn J. E., 1992, ApJS, 78, 1
 Eke V. R., Cole S., Frenk C. S., 1996, MNRAS, 282, 263
 Fernández-Soto A., Lanzetta K. M., Yahil A., 1999, ApJ, 513, 34
 Gioia I. M., Luppino G. A., 1994, ApJS, 94, 583
 Henry J. P., 2000, ApJ, 534, 565
 Hoekstra H., Franx M., Kuijken K., Squires G., 1998, ApJ, 504, 636 (HFKS)
 Hoekstra H., Franx M., Kuijken K., 2000, ApJ, 532, 88 (HFK)
 Jørgensen I., Franx M., Kjaergaard P., 1995, MNRAS, 273, 1097
 Kaiser N., Squires G., 1993, ApJ, 404, 441
 Kaiser N., Squires G., Broadhurst T., 1995, ApJ, 449, 460
 Kelson D. D., van Dokkum P. G., Franx M., Illingworth G. D., Fabricant D., 1997, ApJ, 478, L13
 Luppino G. A., Gioia I. M., 1992, A&A, 265, L9
 Luppino G. A., Kaiser N., 1997, ApJ, 475, 20
 Mellier Y., 1999, ARA&A, 37, 127
 Miralda-Escudé J., 1991, ApJ, 370, 1
 Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
 Schlegel D., Finkbeiner D., Davis M., 1998, ApJ, 500, 525
 Smail I., Ellis R. S., Dressler A., Couch W. J., Oemler A., Jr, Sharples R. M., Butcher H., 1997, ApJ, 479, 70
 Squires G., Kaiser N., 1996, ApJ, 473, 65
 Tyson J. A., Fischer P., 1995, ApJ, 446, L55
 van Dokkum P. G., 1999, PhD thesis, Groningen Univ.
 van Dokkum P. G., Franx M., 1996, MNRAS, 281, 985
 van Dokkum P. G., Franx M., Kelson D. D., Illingworth G., 1998, ApJ, 504, L17
 van Dokkum P. G., Franx M., Kelson D. D., Illingworth G. D., 2001, ApJ, 553, L39
 Voit M. et al., 1997, *HST* Data Handbook, version 3.0. STScI, Baltimore
 Wu X.-P., Chiueh T., Fang L.-Z., Xue Y.-J., 1998, MNRAS, 301, 861

This paper has been typeset from a \LaTeX file prepared by the author.